

Elimination of metabolic pathways to all traditional fermentation products increases ethanol yields in *Clostridium thermocellum*

Beth Papanek^{a,b,c,1}, Ranjita Biswas^{a,b,1,2}, Thomas Rydzak^{a,b}, Adam M Guss^{a,b,c*}

^a Biosciences Division, Oak Ridge National Laboratory, Oak Ridge TN 37830, United States.

^b BioEnergy Science Center, Oak Ridge National Laboratory, Oak Ridge TN 37830, United States.

^c Bredeesen Center for Interdisciplinary Research and Graduate Education, University of Tennessee – Knoxville, Knoxville, TN 37996 USA

¹ These authors contributed equally to this work

² Present address: Centre for Rural Development and Technology, Indian Institute of Technology Delhi, Hauz Khas, New Delhi 110016, India

*Correspondence should be addressed to A.M.G: One Bethel Valley Road, Oak Ridge, TN 37831-6038 USA. Fax: + 1-865-576-8646. Email: gussam@ornl.gov

Notice: This manuscript has been authored by UT-Battelle, LLC, under Contract No. DE-AC0500OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this manuscript, or allow others to do so, for the United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

Abstract

Clostridium thermocellum has the natural ability to convert cellulose to ethanol, making it a promising candidate for consolidated bioprocessing (CBP) of cellulosic biomass to biofuels. To further improve its CBP capabilities, a mutant strain of *C. thermocellum* was constructed (strain AG553; *C. thermocellum* $\Delta hpt \Delta hydG \Delta ldh \Delta pfl \Delta pta-ack$) to increase flux to ethanol by removing side product formation. Strain AG553 showed a two- to threefold increase in ethanol yield relative to the wild type on all substrates tested. On defined medium, strain AG553 exceeded 70% of theoretical ethanol yield on lower loadings of the model crystalline cellulose Avicel, effectively eliminating formate, acetate, and lactate production and reducing H₂ production by fivefold. On 5 g/L Avicel, strain AG553 reached an ethanol yield of 63.5% of the theoretical maximum compared with 19.9% by the wild type, and it showed similar yields on pretreated switchgrass and poplar. The elimination of organic acid production suggested that the strain might be capable of growth under higher substrate loadings in the absence of pH control. Final ethanol titer peaked at 73.4 mM in mutant AG553 on 20 g/L Avicel, at which point the pH decreased to a level that does not allow growth of *C. thermocellum*, likely due to CO₂ accumulation. In comparison, the maximum titer of wild type *C. thermocellum* was 14.1 mM ethanol on 10 g/L Avicel. With the elimination of the metabolic pathways to all traditional fermentation products other than ethanol, AG553 is the best ethanol-yielding CBP strain to date and will serve as a platform strain for further metabolic engineering for the bioconversion of lignocellulosic biomass.

42 **Highlights:**

- 43 • Eliminated genes involved in production of acetate, lactate, formate, and most H₂
- 44 • Highest ethanol yield achieved to date in *C. thermocellum*
- 45 • More than doubled ethanol yield from pretreated switchgrass and poplar
- 46 • Increased ethanol titer at higher cellulose loadings

1. Introduction

There is a strong push in the biofuel industry to move away from traditional starch-based feedstocks that are currently used to produce the majority of the United States' ethanol to more sustainable lignocellulosic crops (Klein-Marcuschamer and Blanch, 2015). This transition could reduce the biofuel industry's dependence on food crops as well as reduce the environmental strain of growing the necessary crops (Brehmer et al., 2008). One enormous challenge in using lignocellulosic feedstocks is the cost-effective deconstruction of recalcitrant plant cell walls to liberate the fermentable sugars (Himmel et al., 2007). Consolidated bioprocessing (CBP) is an approach to overcome this obstacle by using one or more organisms to degrade plant biomass in a single fermentation tank without the addition of commercial enzymes. CBP allows for the simultaneous hydrolysis of cellulose into soluble, metabolizable sugars, and fermentation of the resulting sugars to the product of interest (Lynd et al., 2005). *Clostridium thermocellum*, a thermophilic, anaerobic, cellulolytic bacterium, is one of the most promising microbial candidates for this process. The presence of a large membrane-associated enzyme complex called the cellulosome gives it the ability to hydrolyze lignocellulosic biomass directly to fermentable sugars (Lamed and Bayer, 1988). It is then able to convert the sugars to ethanol and other products such as lactate, acetate, formate, hydrogen, and excreted amino acids such as valine and alanine (Ellis et al., 2012). Further, there is a proven and usable set of tools for genetically engineering the organism (Argyros et al., 2011; Guss et al., 2012; Olson and Lynd, 2012; Tripathi et al., 2010), allowing for the rational manipulation of the *C. thermocellum* genome for improved phenotypes.

In addition to producing ethanol, *Clostridium thermocellum* naturally produces a range of organic acids and hydrogen. The generation of lactate, formate and acetate diverts carbon and electron flux away from ethanol and can quickly make fermentation conditions toxic, presumably due to pH, arresting growth and fermentation. To maximize ethanol yield from *C. thermocellum*, competing pathways must be eliminated. Previously, strains have been engineered to strategically eliminate one or two of these side products, including strains deficient in the production of (i) acetate via deletion of phosphotransacetylase (*pta*) (Tripathi et al., 2010), (ii) lactate via deletion of lactate dehydrogenase (*ldh*) and acetate via *pta* (Argyros et al., 2011), (iii) H₂ via deletion of hydrogenase maturase *hydG* to inactivate all three [FeFe] hydrogenases and *ech* to eliminate the [NiFe] Energy Converting Hydrogenase (Biswas et al., 2015), and (iv) formate via deletion of pyruvate-formate lyase (*pfl*) (Rydzak et al., 2015). Other mutant strains aimed at increasing ethanol production include deletion of lactate dehydrogenase in an ethanol tolerant strain of *C. thermocellum* (Biswas et al., 2014) and deletion of malate dehydrogenase in a strain of *C. thermocellum* that heterologously expressed a pyruvate kinase gene (Deng et al., 2013). However, these strains still produced organic acids and/or H₂ that can acidify the medium and divert carbon and electron flux away from ethanol, reducing ethanol titers and yields. We hypothesized that by combining gene deletions associated with product formation ($\Delta hydG$, Δpfl , $\Delta pta-ack$, and Δldh) into a single strain, side product formation would be effectively eliminated, allowing greater carbon flux through the ethanol production pathway. Therefore, we stacked these mutations into a single strain and characterized the effect on both model substrates and pretreated plant biomass.

2. Materials and Methods

2.1 Growth Media

92 *Escherichia coli* strains were grown on LB medium supplemented with 12 $\mu\text{g ml}^{-1}$
93 chloramphenicol as needed. For *C. thermocellum*, strains were grown either in modified
94 DSM122 rich medium (Tripathi et al., 2010) or defined Medium for Thermophilic Clostridia
95 (MTC) (Linville et al., 2013), with cellobiose, crystalline cellulose (Avicel), or pretreated
96 biomass as the carbon source. To make MTC, Solution A was made in 162 mL serum bottles
97 with cellobiose, Avicel PH105 or dilute-acid pretreated biomass as the carbon source. Dilute
98 acid pretreatment was done previous to experimentation as explained in Schell et al., 2003.
99 Biomass was either poplar or switchgrass previously analyzed with quantitative saccharification
100 from the batch described in (Wilson et al., 2013). The poplar sample was measured to have 646
101 mg glucose equivalents/g biomass $\pm$ 13.6 (other sugar composition included: 19.4 mg xylose
102 equivalents/g biomass $\pm$ 1.6 and 1.6 mg arabinose equivalents/g biomass $\pm$ 0.2). The switchgrass
103 sample was measured to have 522.5 mg glucose equivalents/g biomass $\pm$ 9.3 (other sugar
104 composition included: 72.5 mg xylose equivalents/g biomass $\pm$ 0.4 and 7.1 mg arabinose
105 equivalents/g biomass $\pm$ 1.0). After autoclave sterilization, serum bottles were degassed with
106 filtered N₂ prior to addition of other components. Solutions B, C, E, and F for MTC were filter
107 sterilized through a 0.22 μm filter and solutions D and M were autoclave sterilized. Sterile
108 solutions were mixed in a sterile beaker and 5 mL of the mixed cocktail solution was added to
109 each bottle containing Avicel as the carbon source. For bottles containing cellobiose or biomass
110 as the carbon source, solution M was added to the cocktail mixture and 10 mL of mixed solution
111 was added to each fermentation bottle. For both conditions, the final volume in the serum bottles
112 was 50 mL. The bottles were then degassed with sterile-filtered N₂. When appropriate, *C.*
113 *thermocellum* rich medium was supplemented with 12 $\mu\text{g ml}^{-1}$ thiamphenicol (Tm; Sigma-
114 Aldrich, Saint Louis, MO, USA), 50 $\mu\text{g ml}^{-1}$ 5-fluoro-2'-deoxyuridine (FUdR; Sigma-Aldrich),

or 500 $\mu\text{g ml}^{-1}$ 8-azahypoxanthine (8AZH; Tokyo Chemical Industry, Co., Tokyo, Japan) during strain construction. Plates were solidified with 1.5% agar when appropriate.

2.2 Plasmid and Strain Construction

Plasmid pAMG498 (annotated, Genbank style sequence file, Supplemental File 1) for deletion of *C. thermocellum pta* and *ack* was constructed using yeast gap repair cloning in *Saccharomyces cerevisiae* InvSc1. Standard methods were used to construct *C. thermocellum* deletions (Olson and Lynd, 2012). Briefly, plasmid DNA was isolated from a *dcm*- *E. coli* strain (Guss et al., 2012) and electroporated into *C. thermocellum* strains derived from strain DSM1313 with *hpt* deleted. Electroporated cells were plated on rich medium agar supplemented with Tm. Colonies were picked into liquid rich medium supplemented with Tm, followed by plating dilutions in rich medium supplemented with Tm and FUDR. Colonies were re-streaked on rich medium agar plates supplemented with Tm and FUDR to ensure purity, picked into liquid rich medium supplemented with Tm, re-grown in the absence of Tm, and then plated in rich medium supplemented with 8AZH. Colonies were single colony purified, picked into liquid rich medium, and deletions were confirmed by PCR. Lactate dehydrogenase (Clo1313_1160; *ldh*) was deleted in *C. thermocellum* $\Delta hpt \Delta hydG$ (Biswas et al., 2015) using plasmid pMU1777 (Argyros et al., 2011). Pyruvate-formate lyase (Clo1313_1717; *pflB*) and Pfl-activating enzyme (Clo1313_1716; *pflA*) were deleted in *C. thermocellum* $\Delta hpt \Delta hydG \Delta ldh$ using plasmid pAMG281 (Rydzak et al., 2015). Phosphotransacetylase (Clo1313_1185; *pta*) and acetate kinase (Clo1313_1186; *ack*) were deleted in *C. thermocellum* $\Delta hpt \Delta hydG \Delta ldh \Delta pfl$ using plasmid pAMG498, resulting in *C. thermocellum* strain AG553 (*C. thermocellum* $\Delta hpt \Delta hydG \Delta ldh \Delta pfl \Delta pta\text{-}ack$).

2.3 Fermentation

Strains of *C. thermocellum* were revived from -80°C frozen stocks into MTC medium with 5 g/L cellobiose and grown at 51°C in a Coy anaerobic chamber (Coy Laboratory Products, Grass Lake, MI) until exponential phase (O.D. ~0.4), and 1 mL (2% inoculum) was added to each 162 ml serum bottle, resulting in a working volume of 51 mL. Bottles were incubated at 55°C with 100 rpm orbital shaking. Culture supernatants were sampled every 24 hours and fermentation products were measured by High Performance Liquid Chromatography (HPLC) until production ceased and fermentation product profiles stayed constant. All fermentations were performed twice with three biological replicates both times. The uninoculated MTC medium was found to contain substantial amounts of proline and threonine in addition to the added cysteine, so these three amino acids were removed from the total amino acid analysis in Fig. 2 and 4.

2.4 Analytical Methods

Growth curve data were collected at OD₆₀₀ in a BioTek Eon Microplate reader (BioTek Instruments Inc., Winooski, VT) inside a Coy anaerobic chamber at 55°C.

Fermentation products lactate, formate, acetate, and ethanol were measured using a Breeze HPLC system (Waters, Milford, MA) with an Aminex-HPX-87H column (Bio-Rad, Hercules, CA) and 5 mM sulfuric acid as the mobile phase as previously described (Rydzak et al., 2015). Isobutanol formation was also measured using this method, but none was detected.

H₂ production was measured using an Agilent Technologies 6850 Series II Gas Chromatographer (Agilent Technologies, Santa Clara, CA) with a Carbonex 1010 PLOT (30.0 m x 530 µm I.D.; model Supelco 25467) column as previously described (Yee et al., 2014).

Excreted amino acids were measured using post-column derivatization with ninhydrin using an Aracus Amino Acid Analyzer (membraPure, Berlin, Germany) with a T111 Li-cation exchange column as previously described (Rydzak et al., 2015). All data points used in generating the Figures are included in a supplemental spreadsheet.

3. Results

3.1 Construction of *C. thermocellum* $\Delta hpt \Delta hydG \Delta ldh \Delta pfl \Delta pta-ack$ (strain AG553) mutant

To eliminate H₂ and organic acid fermentation products, *C. thermocellum* strain $\Delta hpt \Delta hydG$ was sequentially modified to delete lactate dehydrogenase (*ldh*), pyruvate-formate lyase (*pfl*), and phosphotransacetylase and acetate kinase (*pta-ack*) resulting in strain AG553, which is predicted to be deficient in the production of lactate, formate, and acetate. This strain also harbors a spontaneous point mutation in the bifunctional aldehyde/alcohol dehydrogenase *adhE* from strain *C. thermocellum* $\Delta hpt \Delta hydG$ (Biswas et al., 2015). The resulting D494G amino acid change expanded cofactor specificity of the alcohol dehydrogenase activity from using only NADH in wild type to using both NADH and NADPH in this strain. Efforts to delete the Ech hydrogenase to eliminate the remaining H₂ production (Biswas et al., 2015) were unsuccessful. Growth was initially tested in minimal medium with cellobiose as the carbon source (Fig. 1). While wild type *C. thermocellum* grew at a rate of $0.14 \pm 0.03 \text{ hr}^{-1}$, strain AG553 experienced a substantially longer lag phase and had an initial growth rate of $0.03 \pm 0.01 \text{ hr}^{-1}$. After reaching approximately an OD of 0.15, the growth rate increased to $0.13 \pm 0.02 \text{ hr}^{-1}$. The observed lag phase was significantly longer for strain AG553. Wild type took 17.5 hours to double its initial OD while AG553 took 39 hours to double its initial OD.

3.2 *C. thermocellum* strain AG553 produces more ethanol on model substrates

Fermentation product formation by *C. thermocellum* AG553 mutant was initially tested on model substrates. When grown in defined medium with 5 g/L of the soluble disaccharide cellobiose as the carbon source, the mutant strain produced greater than two fold more ethanol than the wild type strain with no appreciable amounts of other fermentation products (Fig. 2A). The wild type strain, on the other hand, produced over 15 mM acetate with small amounts of lactate and formate as well. Final ethanol titer reached 32.8 mM for AG553 and 17.7 mM for wild type (56.1% and 30.3% theoretical yield), respectively. Wild type produced 4.2 mM total amino acids, while strain AG553 produced 2.2 mM (Fig. 2A). In both cases, valine was the most abundant amino acid produced (Fig. S1). H₂ production decreased approximately five-fold in strain AG553 relative to the wild type. Total carbon recovery on 5 g/L Cellobiose was 73.5% for wild type and 73.3% for AG553.

On the model crystalline cellulose, Avicel PH105, mutant AG553 again produced three times more ethanol (39.1 mM, 63.5% theoretical yield) than the parent strain (12.3 mM, 19.9% theoretical yield), while the wild type strain produced roughly equal amounts of acetate and ethanol as its primary fermentation products (Fig. 2B). No significant accumulation of organic acids was found in AG553 fermentations. Similar to the cellobiose results, wild type and AG553 produced 4.9 mM and 2.3 mM amino acids, respectively (Fig. 2B), predominantly valine (Fig. S1). Again, H₂ levels decreased five-fold relative to wild type (Fig. 2B). Carbon recovery on 5 g/L Avicel was 80.3% for wild type and 82.3% for AG553.

3.3 *C. thermocellum* strain AG553 produces more ethanol on pretreated biomass

200 Washed, dilute-acid pretreated poplar and switchgrass were used to test the ability of strain
201 AG553 to convert complex plant biomass to ethanol. The amount of available sugars present in
202 this batch of biomass had been previously analyzed, with the pretreated poplar having 646.8 mg
203 glucan/g of dry biomass and the pretreated switchgrass having 522.5 mg glucan/g of dry biomass
204 (Wilson et al., 2013). Therefore, at 5 g/L loading, 3.2 g/L and 2.6 g/L of glucan were available
205 from the poplar and switchgrass, respectively. Control fermentations containing an equal amount
206 of crystalline cellulose (1032 mg glucan/g Avicel) were performed to allow direct comparison
207 for cellulose bioconversion from the pretreated biomass. Therefore, 3.1 g/L and 2.5 g/L Avicel
208 were used as controls for the poplar and switchgrass fermentations, respectively. Culture samples
209 were then taken every 24 hours for HPLC analysis until product formation on biomass no longer
210 increased. For the poplar experiment, fermentation was complete after 288 hours for wild type on
211 poplar, 72 hours for wild type on Avicel, 288 hours for AG553 on poplar, and 96 hours for
212 AG553 on Avicel. For the switchgrass experiment, fermentation was complete after 288 hours
213 for wild type on switchgrass, 24 hours for wild type on Avicel, 288 hours for AG553 on
214 switchgrass, and 96 hours for AG553 on Avicel. Strain AG553 produced 24.9 mM ethanol on
215 Avicel and 23.8 mM ethanol on the pretreated poplar (Fig. 3A), representing 65.5%, and 62.6%
216 respectively, of the theoretical yield of glucan to ethanol. In comparison, wild type *C.*
217 *thermocellum* produced only 13.0 mM ethanol on Avicel (34.2% theoretical yield) and 13.4 mM
218 on poplar (35.3% theoretical yield). Similarly, strain AG553 produced 21.3 mM and 16.6 mM
219 ethanol on Avicel and switchgrass, respectively. These values represent 72.0% and 52.4%
220 conversion of available glucan to ethanol. In comparison, wild type *C. thermocellum* produced
221 11.7 mM and 9.1 mM ethanol on Avicel and switchgrass, respectively (Fig. 3B). This represents
222 39.2% yield of ethanol from the glucan from Avicel and 26% from switchgrass. Organic acid

(lactate, formate, and acetate) production was essentially eliminated in strain AG553 from both poplar and switchgrass. Carbon recovery on poplar was 79.6% for wild type and 71.1% for AG553, while on switchgrass, the carbon recovery was 65.0% and 61.4% for wild type and AG553, respectively.

3.4 Increased ethanol titer in *C. thermocellum* AG553 using higher cellulose loadings

In addition to high yield, high titer ethanol production will also be important for industrial cellulosic biofuel production. We hypothesized that the lack of organic acid production would allow for higher ethanol titers to be achieved. We therefore examined the ability of strain AG553 to convert higher loadings of cellulose to ethanol using Avicel loadings of 1, 5, 10, 20, 30 and 50 g/L in serum bottles with defined medium. Samples were taken every 24 hours until product concentration stopped increasing, which was determined to be 48 hours for wild type *C. thermocellum* at all loadings and 144 hours for strain AG553 at all loadings except for 1 g/L which was 72 hours. Wild type *C. thermocellum* produced approximately equimolar amounts of acetate and ethanol at all loadings (Fig. 4A), and fermentation profiles were similar at all loadings between 5 and 50 g/L, likely due to organic acid accumulation and the resulting pH dropping below 6 (Supp Table #1). With strain AG553, on the other hand, high yields were attained at low cellulose loadings, with 77.9% and 58.0% of theoretical yield at 1 g/L and 5 g/L, respectively. As substrate loading increased, the overall yield decreased based on the total substrate provided (46.6%, 34.0%, 19.7%, and 11% theoretical yield for 10, 20, 30, and 50 g/L, respectively). The titer, on the other hand, reached a maximum of 73.7 mM at a loading of 20 g/L (Fig. 4B), which was not statistically different from the titer at 10, 30, or 50 g/L ($P = 0.07$, 0.10, and 0.08, respectively). All of these values were far below the ~5 g/L ethanol (108 mM) level at which wild type *C. thermocellum* begins to demonstrate slower growth (Herrero and

Gomez, 1980). Interestingly, lactate was produced by strain AG553 at the higher cellulose loadings despite the fact that lactate dehydrogenase has been deleted. Fermentation of sugars to ethanol releases CO₂; therefore, we hypothesized that acidification of the medium from CO₂ accumulation in the serum bottles could also explain the growth and fermentation limitations of strain AG553 at higher cellulose loadings. Therefore, the pH of the fermentation bottles was measured at the end of the experiment, and despite producing far less organic acids, the final pH was substantially lower (Supplemental Table #1) than the optimal pH of 6.8-7.1 for *C. thermocellum* (Garcia-Martinez et al., 1980; Mori, 1990).

4. Discussion

As no known organism is capable of both efficient lignocellulose deconstruction and industrially relevant yield, rate and titer of biofuel production, metabolic engineering is necessary to enable CBP. Previous studies have focused on one or two genetic modifications to increase flux to ethanol, and the capacity for organic acid production had always remained. By deleting the *ldh*, *hydG*, *pfl*, and *pta-ack* genes together, we were able to effectively eliminate lactate, acetate, and formate production and reduce H₂ production on model substrates as well as pretreated poplar and switchgrass. This resulted in a two to three-fold increase in ethanol yields when compared with the wild type strain to a maximum of ca. 70% of theoretical yield and decreased medium acidification. Previously, the highest reported yield was just under 64% theoretical (Biswas et al., 2015) making strain AG553 capable of the highest ethanol yield reported in *C. thermocellum*. By generating a strain of *C. thermocellum* that does not produce substantial amounts of organic acids from cellulosic substrates, we have created a metabolic engineering platform for future engineering efforts to produce fuels and chemicals from lignocellulosic biomass.

268 Despite the elimination of the lactate dehydrogenase protein, when the mutant strain was
269 challenged with higher loadings of crystalline cellulose, it exhibited some minor lactate
270 production at later stages of fermentation (Fig. 4B). It has been previously reported that the
271 malate dehydrogenase protein can exhibit promiscuous behavior producing lactate (Li et al.,
272 2012). This could explain the small and sporadic lactate production observed especially under
273 high cell stress fermentation conditions, and could indicate a metabolic bottleneck that still needs
274 to be addressed. Though side products were removed and ethanol yields were improved, they are
275 still values below those necessary for industrial implementation. Approximately the same
276 amount of secreted amino acids was seen in strain AG553 as in the wild type. This is still a
277 significant sink for carbon and electrons that could otherwise have gone to ethanol production.
278 Future work will need to address excess amino acid production in *C. thermocellum*.

279 Glycolysis in *C. thermocellum* converts glucose to pyruvate and generates two NAD(P)H.
280 Conversion of pyruvate to acetyl-CoA generates two pairs of electrons in the form of reduced
281 ferredoxin (Fd_{red}). The two pairs of electrons on NAD(P)H can be used to reduce one acetyl-CoA
282 to ethanol via AdhE, which can only use NAD(P)H as an electron donor. In this scenario, only
283 50% of theoretical yield could be achieved. In this study, strain AG553 reached yields greater
284 than 50% of theoretical, implying that *C. thermocellum* transfers electrons from Fd_{red} to
285 NAD(P)H that can be used for reduction of acetyl-CoA to ethanol. While electron flux pathways
286 in *C. thermocellum* are not fully elucidated, electron transfer from Fd_{red} to NAD(P)H can be
287 catalyzed by Rnf (Clo1313_0061-0066; $\text{Fd}_{\text{red}} + \text{NAD}^+ + \text{H}^+_{\text{in}} \rightarrow \text{Fd}_{\text{ox}} + \text{NADH} + \text{H}^+_{\text{out}}$) or
288 NfnAB (Clo1313_1848-1849; $\text{Fd}_{\text{red}} + \text{NADH} + 2 \text{NADP}^+ + \text{H}^+_{\text{in}} \rightarrow \text{Fd}_{\text{ox}} + \text{NAD}^+ + 2$
289 NADPH). This additional NADH can be generated by the conversion of reduced ferredoxin
290 (Fd_{red}) by Rnf.

291 *C. thermocellum* strain AG553 produced approximately four-fold less H₂ due to the deletion of
292 the gene encoding the FeFe hydrogenase maturase *hydG*. Previous work combining the *hydG* and
293 the genes encoding the NiFe hydrogenase *ech* resulted in complete elimination of H₂ as a
294 fermentation product (Biswas et al., 2015). Ideally, strain AG553 would be further modified to
295 remove *ech*, such that no H₂ is produced. However, we were not successful in creating this
296 mutation, suggesting the possibility that it is essential under the conditions tested. This might not
297 be surprising, as production of cell biomass results in the production of excess reducing
298 equivalents, and production of a more reduced compound is needed in order to prevent a redox
299 imbalance (Fuhrer and Sauer, 2009). In yeast, for example, this is remedied via glycerol
300 production (Ansell et al., 1997); in *C. thermocellum*, H₂ production likely fulfills this role,

301 especially in strain AG553, which lacks Pfl. Alternate redox-balancing strategies will likely be
302 needed in the future to allow *ech* to be deleted.

303 For industrial ethanol production, it will be important to use inexpensive, renewable feedstocks.
304 Model substrates such as cellobiose and crystalline Avicel are beneficial for the rapid testing of
305 phenotypes in a research setting, but it is important to transition to complex plant biomass
306 feedstocks such as poplar and switchgrass. In this work, strain AG553 had similar ethanol yields
307 on biomass and model substrates such as Avicel, suggesting that the strain is not inhibited by the
308 complexity of real-world biomass feedstocks. This makes the strain an excellent candidate for
309 both continued research on lignocellulosic biofuel production as well as improving the
310 lignocellulosic biofuel production process. Thus, *C. thermocellum* strain AG553 can serve as a
311 platform strain for further organism and process optimization in the future.

312 **5. Conclusion**

313 By combining deletions ($\Delta hydG$, Δpfl , $\Delta pta-ack$, and Δldh) in *C. thermocellum*, we have
314 achieved the highest ethanol yields reported to date on both model substrates and pretreated
315 biomass with *C. thermocellum* strain AG553. The elimination of pathways for organic acid
316 production increased the final ethanol titer achieved at higher cellulose loadings, but further
317 work will be needed to improve ethanol yield and titer when challenged with higher amounts of
318 substrate. Overall, *C. thermocellum* strain AG553 represents a new platform strain for future
319 genetic engineering and process optimization for consolidated bioprocessing of lignocellulose to
320 fuels and chemicals.

322 **Abbreviations**

323 8AZH, 8-azahypoxanthine; CBP, consolidated bioprocessing; FUdR, 5-fluoro-2'-deoxyuridine;
324 Tm, thiamphenicol

325

326 Acknowledgements

327 This work was supported by the BioEnergy Science Center, U.S. DOE Bioenergy Research
328 Center supported by the Office of Biological and Environmental Research in the DOE Office of
329 Science. Oak Ridge National Laboratory is managed by UT-Battelle, LLC, for the U.S. DOE
330 under contract DE-AC05-00OR22725. The funders had no role in study design, data collection
331 and analysis, decision to publish, or preparation of the manuscript.

332

333 References

- 334 Ansell, R., Granath, K., Hohmann, S., Thevelein, J. M., Adler, L., 1997. The two isoenzymes for yeast
335 NAD(+)-dependent glycerol 3-phosphate dehydrogenase encoded by GPD1 and GPD2 have
336 distinct roles in osmoadaptation and redox regulation. *Embo J.* 16, 2179-2187.
- 337 Argyros, D. A., Tripathi, S. A., Barrett, T. F., Rogers, S. R., Feinberg, L. F., Olson, D. G., Foden, J. M., Miller,
338 B. B., Lynd, L. R., Hogsett, D. A., Caiazza, N. C., 2011. High ethanol titers from cellulose using
339 metabolically engineered thermophilic, anaerobic microbes. *Appl Environ Microb.* 77, 8288-
340 8294.
- 341 Biswas, R., Prabhu, S., Lynd, L. R., Guss, A. M., 2014. Increase in Ethanol Yield via Elimination of Lactate
342 Production in an Ethanol-Tolerant Mutant of *Clostridium thermocellum*. *PloS one.* 9, e86389.
- 343 Biswas, R., Zheng, T., Olson, D. G., Lynd, L. R., Guss, A. M., 2015. Elimination of hydrogenase active site a
344 ssembly blocks H₂ production and increases ethanol yield in *Clostridium thermocellum*.
345 *Biotechnol Biofuels.* 8.
- 346 Brehmer, B., Bals, B., Sanders, J., Dale, B., 2008. Improving the corn-ethanol industry: studying protein
347 separation techniques to obtain higher value-added product options for distillers grains.
348 *Biotechnology and bioengineering.* 101, 49-61.
- 349 Deng, Y., Olson, D. G., Zhou, J., Herring, C. D., Joe Shaw, A., Lynd, L. R., 2013. Redirecting carbon flux
350 through exogenous pyruvate kinase to achieve high ethanol yields in *Clostridium thermocellum*.
351 *Metabolic engineering.* 15, 151-8.
- 352 Ellis, L. D., Holwerda, E. K., Hogsett, D., Rogers, S., Shao, X., Tschaplinski, T., Thorne, P., Lynd, L. R., 2012.
353 Closing the carbon balance for fermentation by *Clostridium thermocellum* (ATCC 27405).
354 *Bioresource technology.* 103, 293-9.

Fuhrer, T., Sauer, U., 2009. Different biochemical mechanisms ensure network-wide balancing of reducing equivalents in microbial metabolism. *J Bacteriol.* 191, 2112-21.

Garcia-Martinez, D. V., Shinmyo, A., Madia, A., Demain, A. L., 1980. Studies on Cellulase Production by *Clostridium thermocellum*. *Eur J Appl Microbiol.* 9, 189-197.

Guss, A. M., Olson, D. G., Caiazza, N. C., Lynd, L. R., 2012. Dcm methylation is detrimental to plasmid transformation in *Clostridium thermocellum*. *Biotechnology for biofuels.* 5, 30.

Herrero, A. A., Gomez, R. F., 1980. Development of ethanol tolerance in *Clostridium thermocellum*: effect of growth temperature. *Appl Environ Microb.* 40, 571-7.

Himmel, M. E., Ding, S. Y., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., Foust, T. D., 2007. Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science.* 315, 804-7.

Klein-Marcuschamer, D., Blanch, H. W., 2015. Renewable fuels from biomass: Technical hurdles and economic assessment of biological routes. *AIChE Journal.* 61, 2689-2701.

Lamed, R., Bayer, E. A., 1988. The Cellulosome of *Clostridium thermocellum*. *Adv Appl Microbiol.* 33, 1-46.

Li, Y., Tschaplinski, T. J., Engle, N. L., Hamilton, C. Y., Rodriguez, M., Jr., Liao, J. C., Schadt, C. W., Guss, A. M., Yang, Y., Graham, D. E., 2012. Combined inactivation of the *Clostridium cellulolyticum* lactate and malate dehydrogenase genes substantially increases ethanol yield from cellulose and switchgrass fermentations. *Biotechnology for biofuels.* 5, 2.

Linville, J. L., Rodriguez, M., Jr., Land, M., Syed, M. H., Engle, N. L., Tschaplinski, T. J., Mielenz, J. R., Cox, C. D., 2013. Industrial robustness: understanding the mechanism of tolerance for the *Populus* hydrolysate-tolerant mutant strain of *Clostridium thermocellum*. *PloS one.* 8, e78829.

Lynd, L. R., van Zyl, W. H., McBride, J. E., Laser, M., 2005. Consolidated bioprocessing of cellulosic biomass: an update. *Current opinion in biotechnology.* 16, 577-83.

Mori, Y., 1990. Characterization of a Symbiotic Coculture of *Clostridium thermohydrosulfuricum* Ym3 and *Clostridium thermocellum* Ym4. *Appl Environ Microb.* 56, 37-42.

Olson, D. G., Lynd, L. R., 2012. Transformation of *Clostridium thermocellum* by electroporation. *Method Enzymol.* 510, 317-30.

Rydzak, T., Lynd, L., Guss, A., 2015. Elimination of formate production in *Clostridium thermocellum*. *J Ind Microbiol Biot.*

Schell, D. J., Farmer, J., Newman, M., McMillan, J. D., 2003. Dilute-sulfuric acid pretreatment of corn stover in pilot-scale reactor - Investigation of yields, kinetics, and enzymatic digestibilities of solids. *Appl Biochem Biotech.* 105, 69-85.

Tripathi, S. A., Olson, D. G., Argyros, D. A., Miller, B. B., Barrett, T. F., Murphy, D. M., McCool, J. D., Warner, A. K., Rajgarhia, V. B., Lynd, L. R., Hogsett, D. A., Caiazza, N. C., 2010. Development of *pyrF*-based genetic system for targeted gene deletion in *Clostridium thermocellum* and creation of a *pta* mutant. *Appl Environ Microb.* 76, 6591-9.

Wilson, C. M., Rodriguez, M., Jr., Johnson, C. M., Martin, S. L., Chu, T. M., Wolfinger, R. D., Hauser, L. J., Land, M. L., Klingeman, D. M., Syed, M. H., Ragauskas, A. J., Tschaplinski, T. J., Mielenz, J. R., Brown, S. D., 2013. Global transcriptome analysis of *Clostridium thermocellum* ATCC 27405 during growth on dilute acid pretreated *Populus* and switchgrass. *Biotechnol Biofuels.* 6, 179.

Yee, K. L., Rodriguez, M., Jr., Thompson, O. A., Fu, C., Wang, Z. Y., Davison, B. H., Mielenz, J. R., 2014. Consolidated bioprocessing of transgenic switchgrass by an engineered and evolved *Clostridium thermocellum* strain. *Biotechnology for biofuels.* 7, 75.

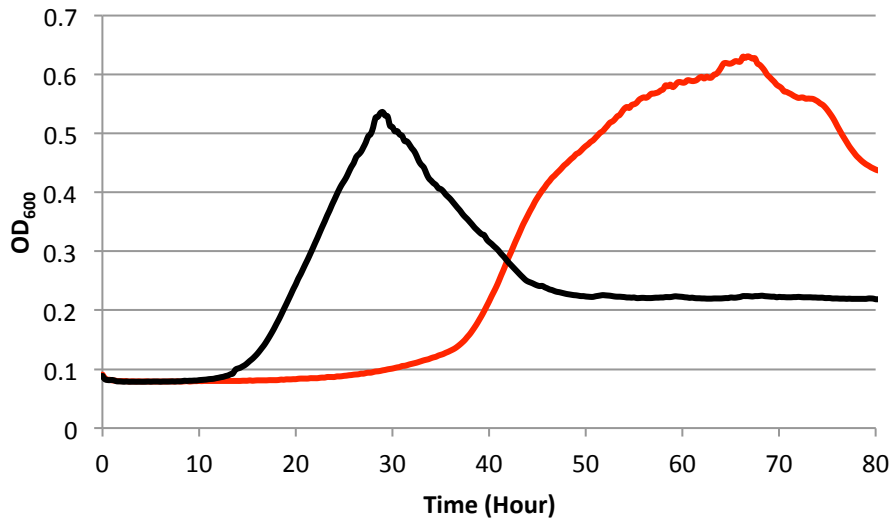


Fig. 1. Growth curve of *C. thermocellum* mutant in defined media with 5 g/L cellobiose. The growth data was collected on a plate reader with measurements taken every 15 minutes. Lines are the average of triplicate cultures. Black line, wild type; Red line, *C. thermocellum* strain AG553.

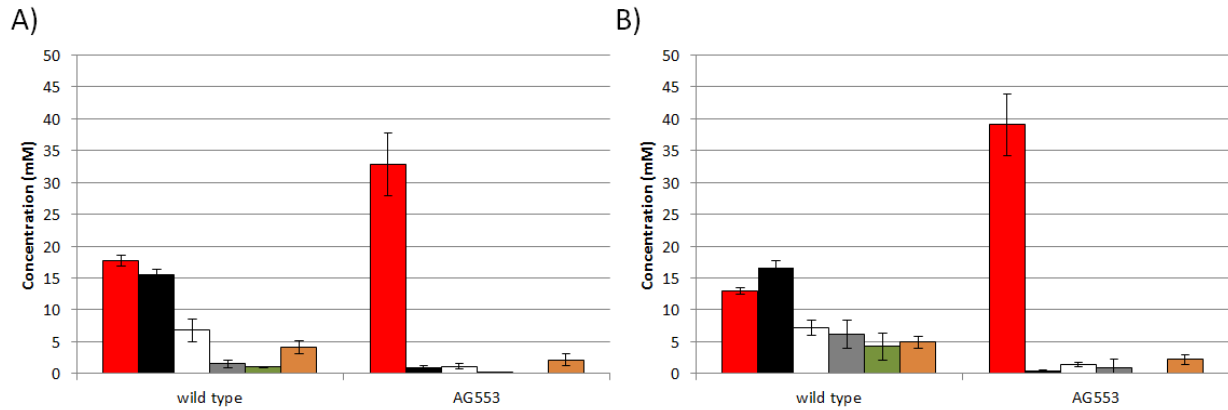


Fig. 2. *C. thermocellum* mutant product formation on model substrates. Growth on 5 g/L (A) cellobiose and (B) crystalline cellulose Avicel in defined MTC media. Red, Ethanol; Black, Acetate; White, H₂; Grey, Lactate; Green, Formate; Orange, Excreted Amino Acids. H₂ concentration is reported as mmol H₂ in the gas phase per L liquid medium to facilitate comparisons.

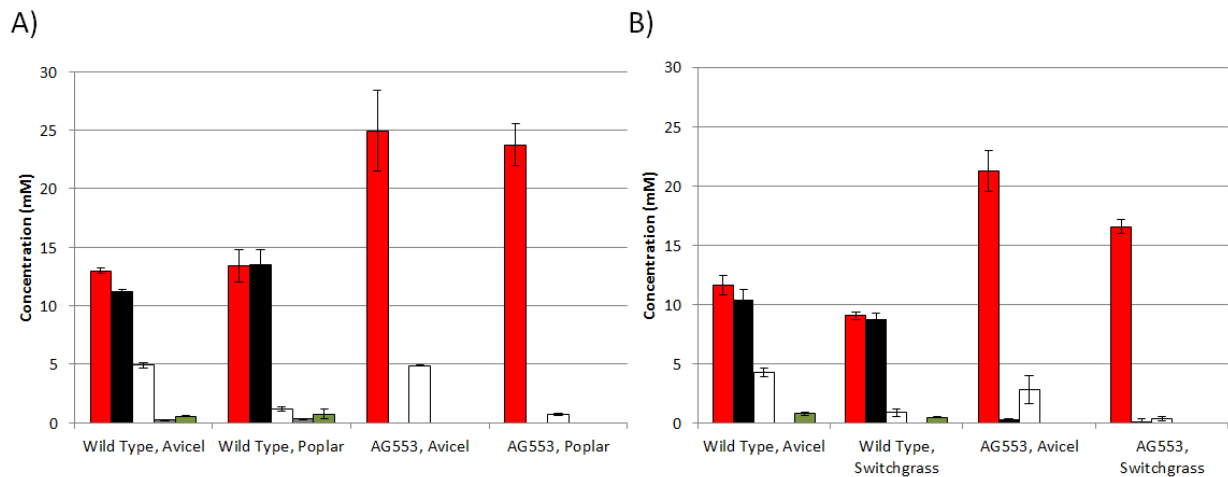


Fig. 3. *C. thermocellum* mutant product formation on pretreated plant biomass. Avicel samples taken after 96 hours of fermentation and pretreated biomass samples taken after 288 hours of fermentation for both wild type *C. thermocellum* and *C. thermocellum* strain AG553 (A) Fermentation products from 3.2 g/L Avicel and 5 g/L washed, dilute acid pretreated poplar and (B) 2.6 g/L Avicel and 5 g/L washed, dilute acid pretreated switchgrass. Red, Ethanol; Black, Acetate; White, H₂; Grey, Lactate; Green, Formate.

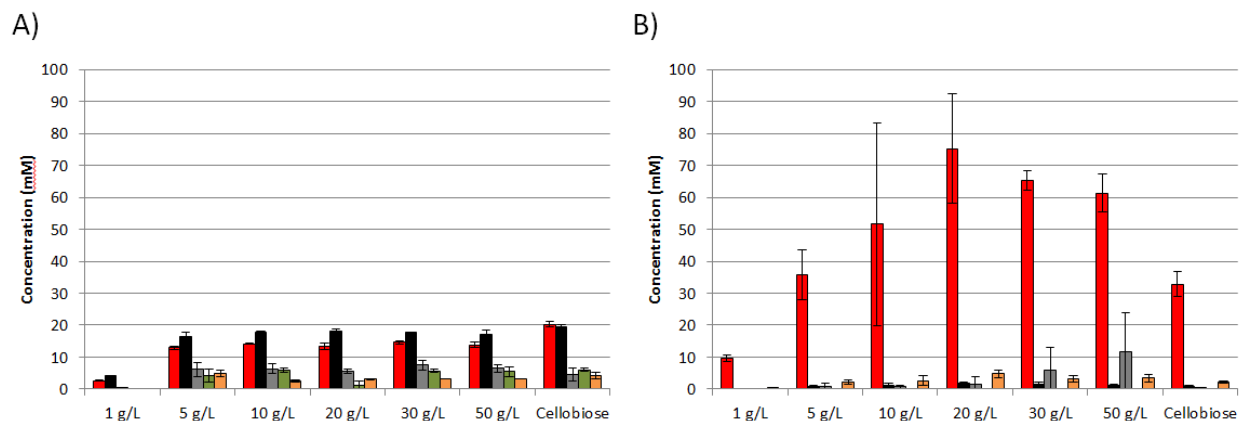
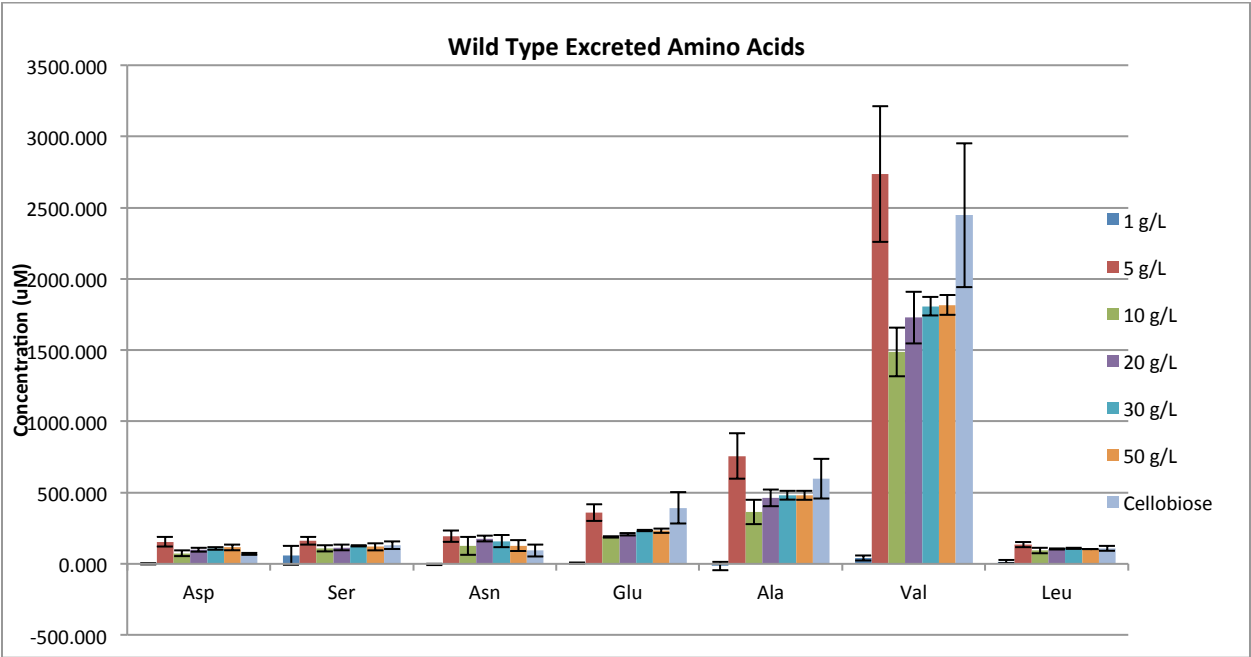
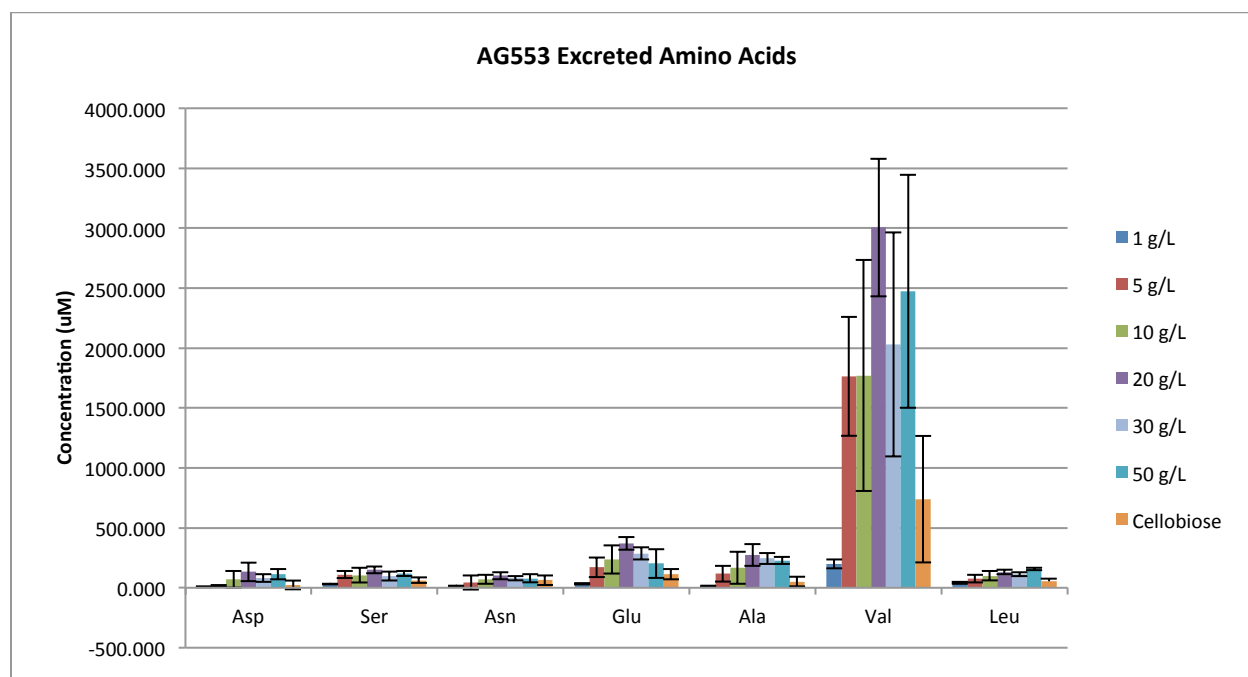


Fig. 4. *C. thermocellum* mutant product formation on different loadings of Avicel PH105 in defined MTC medium. (A) Wild Type *C. thermocellum* (B) *C. thermocellum* strain AG553.
 Red, Ethanol; Black, Acetate; Grey, Lactate; Green, Formate; Orange, Excreted Amino Acids.

437 A)



439 B)



440

441 **Fig. S1. Concentrations of excreted amino acids produced at various substrate loadings. A)**

442 **Wild Type and B) strain AG553**

443

444 Table S1

Substrate Concentration	Δhpt , final pH	AG553, final pH
0 g/L Avicel	7.99 ± 0.22	7.53 ± 0.02
1 g/L Avicel	7.26 ± 0.19	7.11 ± 0.02
5 g/L Avicel	6.13 ± 0.30	6.57 ± 0.09
10 g/L Avicel	6.07 ± 0.31	5.45 ± 0.15
20 g/L Avicel	6.07 ± 0.28	5.30 ± 0.05
30 g/L Avicel	6.06 ± 0.29	5.60 ± 0.42
50 g/L Avicel	6.11 ± 0.34	5.72 ± 0.19
5 g/L Cellobiose	6.19 ± 0.26	6.37 ± 0.02

445